Indian Journal of Science and Technology, Vol 9(36), DOI: 10.17485/ijst/2016/v9i36/102911, September 2016
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Physical Design Implementation of 16 Bit Risc Processor

Devaraconda Dinesh and R. Manoj Kumar

School of Electronics Engineering, VIT University, Chennai-600127, Tamil Nadu, India;
devaraconda.dinesh2014@vit.ac.in, manojkumar.r@vit.ac.in

Abstract

Objectives: To design the 16bit Reduced Instruction Set Computing (RISC) processor using the Verilog Hardware Description Language (HDL). Methods: This is a 4 stage pipelined processor with idle state, fetch state, decode state, and execute state. Here the write back stage is also performed in the same execution state. The physical design of the processor is done by floor plan followed by placement and routing process. In the placement process iteration is continued till it meet the timing constraints. **Findings:** In the floor plan we decide the utilization factor, width and height of the core and die, apart from that we decide the location of the IO Pads. The location of the preplaced cells and standard cell placement are find during the placement process. In the placement process if the congestion is huge then the hard blockage placement is done to eliminate the congestion. During the trail route the tool does the timing analysis using the ideal clock, which is then replaces the ideal clock with the real clock after the clock tree synthesis. This process is followed by the spare cells placement in the empty region where we can replace these cells with the logic cells in future; this step in the physical design procedure is known as Engineering Change Order (ECO). During the tape out process if the timing is not met we may not do the placement and routing process from the starting step so we go for the ECO where we can replace the cells with the other cells of same functionality to reduce the delay in order to remove the setup slack. **Applications:** RISC processors are used in wide range of applications such as mobile phone, tablet computer, super computers etc, RISC processors are also used in signal processing applications such as convolution, correlation etc.

Keywords: Engineering Change Order (ECO), Pipelined Architecture, RISC Processor

1. Introduction

Reduced instruction set computer is a processor with reduced and less complexity of instructions. Some of the commonly used processors are SPARC IBM's PowerPC and MIPS. As the design of the CISC processor became more complex and did not met the expected criteria people shifted for other alternatives. It has been observed that the speed of the processor gets reduced when the memory and the processor are interacted. Thus the Instruction set architecture of the processor is made simple. As a result in most of the RISC processors only few instructions will be used for the interaction of the memory and the processor. The instructions complexity in the CISC processor is reduced in the RISC processor by the usage of opcode in the instruction itself. This is a four stage pipelined architecture which are the fetch stage decode stage execute stage and the write back stage. Initially the data is fetch and it is sent for the decode and the execution state during the decode state the processor will be fetching the next data instead of being in the idle state. During the write back state the processor will be fetching the data followed by decode and execution state.

2. Description of Signals

This work shows the implementation of the 16bit reduced instruction set computer processor with each instruction is of 16 bits and there are 16 registers also clock is used for the data path and control path. Different approaches have been done to increase the execution rate. Some of the fre-

*Author for correspondence

quently used ways are prefetching and pipelining.When the present instruction is fetched then the next instruction is fetched by using the instruction queue before the execution of the previous instruction. In case of the 8086 processor which is a 16 bit processor the instruction is fetched in to the queue which is of 6 bytes and the second byte of the queue is used to store the next instruction which makes the design more parallel and efficient.

3. Instruction Set (Is)

The IS consists of 16 instructions. The size of each instruction is 16 bits wide, which is sub divided in to 4 sections of opcode, destination register, source register and target register. Each opcode refers to the particular operation being performed by the processor.Consider a case where when the opcode is 0000 then the operation being performed is the addition operation where the content of the destination register and the target register are added up and stored in the source register. Similarly when the opcode is of 0001 subtraction operation is performed; here all the arithmetic, logical, bitwise operations are performed by the reduced instruction set computer processor.

4. Control Unit Design

The control unit has four different states which represents the pipelining process. IDEL, FETCH, DECODE and EXECUTE stage. Here the Fetch state is sub divided in to operand state and the instruction state. Whenever the reset signal is asserted then FSM is made to IDEL state. During the IDEL state it makes the pc_count signal to 1 to make it as the current program count. When the reset signal is de asserted then the fsm enters in to the fetch stage to take the data from the memory address of location 0 is loaded in IR for the execution of the program. Since the next state is the fetch state asserts the ir_write signal and selects both the operands and performs the addition operation.

Each and every signal transition happens during the first rising edge of the fetch stage. ALU operand selection, operand b selection and the data selection lines are the one hot signal for decoding the data path. The next operation is determined by the opcode which is obtained from the instruction register, the program counter which is incremented is set from the output of ALU[1]. The data values are loaded from the both the operands from the register. The load operation will need only one operand while the store operation needs the two operands, one is the data and the other is the address word. The instructions of the branch operation uses the offset of the in its instruction and the incremented program counter is an operand to the ALU. The jump instruction stores the PC + 1 while the jump return loads the address to be returned in to the operand. After the fetch state it enters in to the execute state. In the first stage, of ALU operation all the control signals are sent from the control unit to the ALU. When the program counter is to be incremented then the control unit will assert the signal pc_sel to select the pc signal and asserts the pc_write, ope_sel selects the program counterand the opbb_sel selects the 10 case which chooses the 1 and thalu_sel is set with the 001 case where it selects the both the operands and does the addition operation. Thus the program counter is incremented.

5. Physical Design Implementation

Physical design is a way of converting the logical expressions in to layout**.** The complete cycle is divided in to several steps like partitioning, floorplanning, placement, clock tree synthesis, and routing, static timing analysis till the tape out process.During the partition process different algorithms are followed to partition the blocks based on the size of the blocks and the interconnection between them. Generally the net list describes the interconnection between the blocks. The partition is done in the hierarchical level where the sub module are done first followed by the top most module.

6. I/Opads

Generally the chip has to interact with the external world components like the audio or video devices but there can be a large capacitance connected through the chip. So considering these effects a circuitry is designed to connect the chip with the external devices. This problem is removed by the usage of the I/O pads; the pins of the chip are connected to the pads with the help of bonding wires. There are two types of power pads one is the core pad and the other is the power pad. VDD and VSS are the power pads which are connected to the design. At least there should be one vdd and vss in the design to get the power supply and the ground connection to the chip. Core pads are of both the input pads as well as the output pads and

also there is other kind of pad which is the bidirectional pads apart from that we have the special pads[2]. The special pads include the filler pads which are used to mainly fill the gaps between the IO pads. Bidirectional pads act as both the input pads as well as the output pads the operation of the pads is determined by the control signal. When the input control signal is 1 the bidirectional pad may act as the input pad and when the control signal is 0 the bidirectional pad may act as the output pad.

7. Floorplan

Floor planning of the design is done to give a blueprint of where the location of the macros and module has to be placed in order to increase performance the location of the modules is based on the total area of the module and macros are found out to minimize the area. The inputs for the floorplan are the net list, timing libraries, and the library exchange format files i.e. LEF files. Floorplan includes the black box and pins placement power planning which includes both the adding power rings and also adding power stripes to the design. By doing the automatic floorplan the tool automatically decides the location of the blocks. But it is generally preferred to do the manual floorplanning for occupying minimal area. Generally the algorithm for the floorplan runs so many iterations to get the optimal result[3]. The performance, timing criteria is taken care by the designer during the floorplan.

The reason we go for the power planning is to avoid the two disadvantages like the power drop and ground bounce concepts.

Power Droop: - Consider an inverter cell of 16 bit bus is used in the design with the load capacitor, so whenever the output of the inverter makes a transition from the logic 0 to logic 1 then all the 16 capacitors will try to charge up to logic 1 voltage at ones due to that because only one vdd source and each capacitor is taking the power from the same source causes the power drop i.e. logic 1 voltage of vdd reduces if the new voltage is within the noise margin range the out may not affect if the range is out of the noise margin the output gets affected, in order to avoid this effect we go for the multiple vdd and multiple vss in the design.

8. Placement

This is a process of fixing the location of the standard cells. We can also have the preplaced cells in our design which means before the floorplan we can fix cells location. The performance of the design gets degraded when results in occupying more area. In the first iteration cells are placed and the trial routing is done to check the Static Timing Analysis (STA) weather it is meeting the timing constraints or not[4]. If not the second iteration is done to reduce the area and for STA. The iteration process continues till the layout occupies the minimum area and meets the specifications of the design.

9. Special Cells

Different kinds of cells are used in the physical design flow for various functionality. After the placement of the standard cells the rest of the area is free so in those areas the spare cells are used to fill the empty spaces. During the tape out scenarios if any cell has to be placed in the design then the spare cell is replaced with the logical cell5 .The usage of the decoupling capacitors is that When a single power supply source is used in the design, considering the resistance of the wire there is a drop of voltage in the wire so instead of voltage vdd reaching the circuit due to the drop in the wire the voltage may be reduced to vdd' if this range is within the noise margin range the output will be perfect so in order to get the exact vdd voltage to the circuit we use the decoupling capacitors. Initially the capacitors are charged to the voltage vdd and the power supply for the design is used from these decoupling capacitors. There may be a case of failure of the functionality because of IR drop. The drop happens when there is a huge current demand in the circuit from the power rings. When the power supply is far then there might be case of causing the metastability. During high current requirement then the capacitors gives its stored charge to the power rings.

10. Clock Tree Synthesis

Duty Cycle: - Consider a clock signal of 1ns where the duty cycle says for what amount of time the clock signal will be high and for what amount of time the clock signal will be in the low state that is calculated in terms of percentage.

In above example for 500ps the clock pulse remains high and for the 500ps the clock pulse remains low here the rise time and the fall time remains same. This implies that the rise time and the fall time is of 50% duty cycle this is expected to be maintained at the launch flop and

the capture flop. This is the first check which is the duty cycle check. When this clock signal is sent through the buffers it might get affected and finds rise time as 58% and fall time as 42%. The reason for this is that there are two kinds of buffers one is the regular buffer and the other is the clock buffer. Consider another case where the clock remains high for the 200ps and remains low for 800ps. In this case the duty cycle is 20% and 80% when it reaches to the flops it might get affected and becomes 15% rise time and 85% fall time.

Latency Check: - Basically we want the latency to be low let L1 be the latency which moves from clock to the starting point of the launch flop clock. The value of latency L1 is sum of 3buffer delays and 4wiredelays. L2 is the sum of the 2buffer delays and 3 wire delays. Low latency implies that the number of buffers required in the clock path reduces which directly affect the power in the chip which is an advantage. By reducing the at least one buffer in each flop path in the million circuit flop then the number of buffers reduces which internally reduces the area. Less latency also leads to the less number of routings and less power consumption.

Clock Tree Power: - CTP is a function of latency as well as the transition[6]. Consider the buffer which is nothing but the back to back connection of the inverter. Consider the characteristics of the inverter, whenever the input is logic 0 then the pmos gets turned on and the nmos gets turned off as a result the capacitor gets charged from the power supply vdd. When the input is greater than the threshold voltage of the nmos, then a small amount of the current flows from the drain to source of the nmos. But ideally no current flows from the drain to source in practical small current flows from the nmos this is known as the sub threshold current this leads to the concept of the leakage current. By considering the sub threshold current for each and every device it counts to a significant number.

Signal Integrity and Crosstalk: - Clock tree in generally built of both buffers and the inverters.

Consider the victim and the aggressor net as shown below. When these two wires are placed too closed together the coupling capacitance comes in to the effect.

$$Capacitance = \frac{€ * A}{D}$$

If the coupling capacitance is greater than the sum of the capacitances at the victim node then the Cm will be having the more dominant effect, which leads to cross talk.

11. Factore Affecting Clock Path

The entire circuit is modeled as the clock port and the lumped resistance and the capacitances and the load capacitances which is the capacitances of the flops.

L=length of the wire
R0=per unit length resistance
C0 is the per unit capacitance.
R0L=Total resistance
C0L=Total capacitance.
RD= Driver resistances.

There is also some resistance offered by the clock port which is known as the RD. So now it becomes simple to analyze this network and do the clock tree buffering.

Time constant = (RD + R0L) * (C0L + Cl)
= RDCl + (RDC0 + R0Cl) L + R0C0 L^2
= O (L^2)

The above function is dominant of the L^2 it means the time required to charge the output load capacitance with the particular length of the wire is proportional toL^2 . Greater the length then more amount of time for the output load capacitance to charge then proportional to the square of the length of the wire. So this is the reason the transitions are having worse effect. Let us try to reduce the length in order to reduce the transition when the buffer 3 is placed as shown its length,

$$=L/4 + L/4 + L/4 + L/4 = L(approx$$

$\frac{L}{4}$. By considering the second buffer the effective length is The load that is seen by the buffer is pretty huge. So at this particular node it is having four fanout which means that it is driving four devices (flip-flops). c1,c2,c3,c4 are the wire capacitances from the wire load model the value of the capacitances is 25fF so at the node of the 3rd buffer we can see the effective capacitance as,

$$C_{total_A} = c1 + c2 + c3 + c4$$

$$C_{total_A} = 25 + 25 + 25 + 25$$

$$C_{total_A} = 100fF$$

100fF is so huge in case of general scenarios the node capacitance at the clock should be in the range of 50-60fF. One of the solutions to reduce the capacitance is to build the buffer tree[7].There are two levels of buffering. The advantage of above case is that the signal sent at the clock pin reaches the flop pin exactly. But there is a penalty over here the extra buffers which are placed will occupy extra area on the chip. Whenever the input is passed through

the buffer it's known as the level1 and similarly when it passes through the other buffers it's called as level2. We observe that there are 2 levels of buffering in the clock tree. Each level can have n number of buffers but the level will be same for all those buffers.

As given in the Figure 1, H – Tree diagram, we have the same values of capacitance for both the nodes A and B which is an advantage. If there is a condition like node A and node B are having different capacitance the one buffer will be having the large delay and the other will be having the small delay due to that the concept of skew would have raised.

At the every level the buffers which are used should be identical the advantage is both the buffers will shows the similar characteristics. Each flop will be driven by one buffer this concept is raised from the load splitting. First we try to find the levels of buffering in the clock tree.

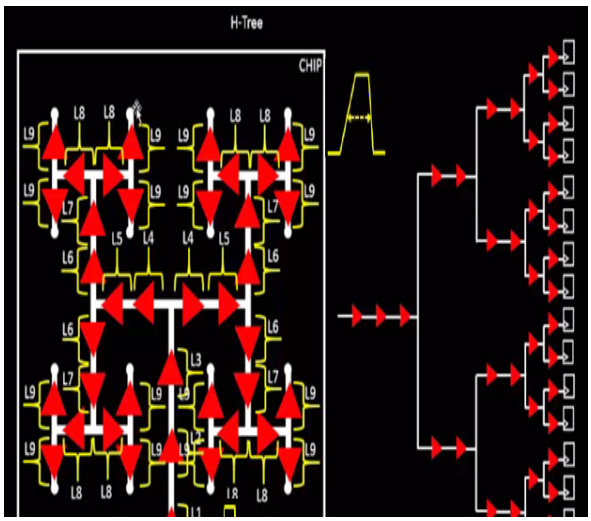

Figure 1. H Tree.

So totally there are 9 levels of buffering and finally at the output of these buffers the flops are connected. The white spots are nothing but the clock pins of the flops. Consider for the 16 flops we got the 9 levels of buffering and for the case where are millions of gates then the number of levels will be more this is the basic concept behind the clock tree this is done for the multimillion gates by doing this algorithm automatically[8]. If we try to understand with the 16 flops the more or less same thing follows for the multimillion flops. The next observation is at each node it should be driving the same load and the buffer should be identical, at level 4 it is clear that the length of the wire is same and the buffers are identical so the capacitances are also the same which is satisfying the criteria of observation. In this particular case we considered the all buffers are identical and also the length of the wires in all the levels is same.

Consider the pulse width check in the clock tree we try to send the rise clock pulse and the fall clock pulse at the input of the clock pin and try to check the pulse width at the clock pin of the flop to ensure the pulse width quality check. As we move the signal from the clock pin we see that the transition gets slightly increased through the buffers. Due to the interference of the wire the transition gets increased so by the time it reaches the clock pin of the flop we see a average detritions of the signal. The same case is for the falling edge of the clock but with the variation less compared to the rising edge. At the input the duty cycle is 50% so the rise time and the fall time are 50% with 500ps of rise time and 500ps of fall time and at the output the rise time gets increased to 700ps and 700ps for fall time and the frequency becomes 800 MHz there is a huge loss of the frequency so it affects the speed of the circuit. The first observation is that the rise transition and the fall transition not the same[9]. Basically the properties of the buffers are like the performance of the buffer which means that the speed of the buffer at what time the output is getting propagated and the second property is that it's symmetry which means that the rise time and the fall time should be same which in this case it is not so if we classify the buffers in terms of symmetry there are two types one is the normal or regular buffer and the clock buffers. The buffers which are used in the clock path are called as the clock buffers and the buffers used in the data path are called as the regular buffers. Data path is nothing but the path between the flops regular buffers shows the property of performance while the clock buffers shows the property of symmetry. Consider the buffer which is nothing but the back to back connected inverters. The main reason behind the variation in the rise and the fall time is that the on resistance of the PMOS is greater than the NMOS on resistance

12. Routing

There are two types of routing one is global routing and the other is the detailed routing. In case of the global routing the interconnections between the standard cells connects loosely, while in case of the detailed routing the detouring of connections are reduced. During the

placement of the standard cells also trial route is done to estimate the static timing analysis weather the design met the timing constraints[10]. The timing analysis is done using the ideal clock which means that the delays of the wires are neglected but after the clock tree synthesis the real clocks are considered which means the buffer delays as well as the wire delays are considered for the analysis. Routing algorithm is done with much iteration to reduce the detouring of the metal wires.

As we can infer from Figure 2, the placement that the red color represents the congestion which is nothing but the overlapping of the standard cells as we discussed earlier we go for the soft, partial, or the hard block placement. Congestion shows a/b routings which means that is the required number of routing tracks while b is the available number of routing tracks. Whenever the numbers of available tracks are less than the number of required tracks then the congestion will occur.

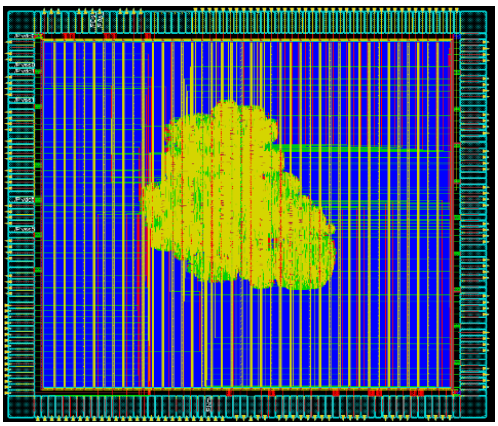

Figure 2. Routing.

13. Static Timing Analysiss

At the time of tape out if we encounter any problem like mismatch in timing then to add any extra cell instead of adding the cell and followed by the floor planning placement and routing the process becomes more complex again so in order to reduce this process we go for the engineering change order where we can add the cell even at the time of the tape out process.

Since the setup slack has occurred so in order to remove the setup slack one of the ways is to reduce the delay of the cell and the other way is to add the buffer in the clock path. In the above example we see that in the data path buf3 has the highest delay because of that the setup slack has occurred so this cell is replaced with the buf1 of the same functionality of lesser delay thus reducing the combinational delay and removing the negative slack. The other way is to add the buffers in the clock path. The above figure represents the positive slack report where the all the failing paths have been removed by doing the engineering change order. As shown in the above timing path analyzer table the worst negative slack is 0 and the total negative slack is 0 the design rule violation also clears with no violations.

14. Conclusion

The physical design implementation of the RISC processor is done by achieving the each and every quality check during the clock tree synthesis to meet the skew, duty cycle, latency, pulse width and the clock tree power compared to the statistical analysis[11]. An Engineering Change Order (ECO) is done to ensure the design to meet the required specifications and timing.

15. References

1. Neeraj Jain. VLSI design and optimized implementation of a risc processor using xilinx tool. International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE). 2012 Dec; 1(10):52–6.
2. Preetam B. Hari Krishna Moorthy. FPGA Implementation of low power pipelined 32-bit RISC Processor. International Journal of Innovative Technology and Exploring Engineering (IJITEE). 2012 Aug; 4(4):1–7.
3. Ibrahim D.16-bit microprogrammable microcomputer with writable control store". IEEE Trans. Computers. 2011 Nov; 39(11):1385–90.
4. Kahng A, Lienig J, Markov I, Hu J. VLSI physical design: From graph partitioning to timing closure. Springer, 2011, 27.
5. Mehrotra, Alok, Van Ginneken, Lukas P, Trivedi, Yatin. Design flow and methodology for 50M gate ASIC, IEEE Conference Publications, 2003.
6. Sharda PK, Jain GP. Design and implementation of 5 Stages pipelined architecture in 32 Bit RISC Processor. International Journal of Emerging Technology and Advanced Engineering. IJETAE, 2012 April; 2(4):340–6.
7. Chia-Ming Chang, Shih-Hsu Huang, Yuan-Kai Ho, Jia-Zong Lin, Hsin-Po Wang, Yu-Sheng Lu. Type-matching clock tree for zero skew clock gating. Design Automation Conference 45th ACM/IEEE, DAC 2008.
8. Wendel D , "POWER7, a highly parallel, scalable multi-core high end server processor". IEEE Journal Solid-State Circuits. 2011; 46(1). p. 145.

9. Reaz MBI, Amin N , Ibrahimy MI , Mohd-Yasin F, Mohammad A. Zero skew clock routing for fast clock tree generation. Canadian Conference on Electrical and Computer Engineering. (CCECE). 2008 May; 4–7:23–8.
10. Tsai JL, Chen TH and Chen CCP. Zeroskewclock tree optimization with buffer insertion/sizing and wire sizing. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems.* April 2004; 23(4):565–72.
11. KV Rao, A Angeline, VSK Bhaaskaran. Design of a 16 bit RISC processor. Indian Journal of Science and Technology. August 2015; 8(20):1–7.